



Ressources naturelles  
Canada

Natural Resources  
Canada



# L'INFRASTRUCTURE CANADIENNE DE DONNÉES GÉOSPATIALES PRODUIT D'INFORMATION 42f

## **Document d'information sur l'interopérabilité géosémantique**

GéoConnexions  
Hickling Arthurs Low Corporation

2015

© Sa Majesté la Reine du chef du Canada, représentée par le ministre de  
Ressources naturelles Canada, 2015

Canada

---

## REMERCIEMENTS

---

GéoConnexions tient à exprimer ses remerciements pour l'assistance offerte par Intelli<sup>3</sup> – Dr Yvan Bédard et Mme Sonia Rivest, lors de la recherche, de la rédaction et de la modification du document, Hickling Arthurs et Low Corporation – M. Ed Kennedy, pour le suivi du processus de révision et de production. Nous aimerions également mentionner l'assistance offerte par Dr Glen Hart, chercheur scientifique principal et directeur de la recherche, Ordnance Survey (Organisation nationale de cartographie de la Grande-Bretagne) et de Dr John Goodwin, scientifique principal, pour avoir partagé leur expérience novatrice en matière d'interopérabilité géosémantique. Le personnel de GéoConnexions a aussi travaillé à la gestion, aux commentaires et à l'orientation de ce projet : Dr Jean Brodeur, Mme Mélodie Bujold, M. Simon Riopel, Mme Cindy Mitchell et Mme Paula McLeod. Par ailleurs, M. Craig Eby et M. Alexandre Cyr ont apporté leur soutien lors du montage du présent document.

# Table des matières

<b>Remerciements .....</b>	<b>i</b>
<b>1. Préambule.....</b>	<b>1</b>
<b>2. Présentation.....</b>	<b>Error! Bookmark not defined.</b>
2.1 Origines de l'interopérabilité géosémantique .....	1
2.2 L'interopérabilité géosémantique dans la société d'aujourd'hui .....	2
<b>3. Définitions de base de l'interopérabilité géosémantique.....</b>	<b>2</b>
3.1 Quels sont les « concepts » et comment les communiquer à des fins d'interopérabilité?.....	3
3.2 Sources classiques de communication de la signification .....	3
3.3 Qu'est-ce que la « sémantique » et quel est son rôle dans un environnement Web? .....	4
3.4 Qu'est-ce qu'une « ontologie formelle », quels sont ses rôles et ses langages? .....	5
3.5 Qu'est-ce que l'« interopérabilité »? .....	5
3.6 Qu'est-ce que le « contexte »? .....	6
3.7 Qu'est-ce que la « géosémantique »? .....	6
3.8 Qu'est-ce que l'« interopérabilité géosémantique »? .....	7
3.9 Langages de l'interopérabilité géosémantique pour le Web.....	7
3.10 Normes de soutien de l'interopérabilité géosémantique.....	8
<b>4. Étude de cas : Ordnance Survey .....</b>	<b>8</b>
<b>5. Défis.....</b>	<b>9</b>
<b>6. Conclusion .....</b>	<b>10</b>
<b>Annexe 1 : Liste d'acronymes .....</b>	<b>11</b>
<b>Annexe 2 : Références.....</b>	<b>12</b>

# 1. Préambule

*Le présent guide fait partie d'une série de documents sur les politiques opérationnelles préparés par GéoConnexions. Il a pour but d'informer les intervenants de l'ICDG sur les principaux concepts et l'importance de l'interopérabilité géosémantique de même que sur les technologies connexes et les exemples d'intégration de l'interopérabilité géosémantique dans les organisations.*

L'interopérabilité géosémantique est une notion fondamentale pour les applications exploitant les données et les services géospatiaux du Web. Néanmoins, elle est souvent perçue comme un concept abstrait difficile à appréhender. Le présent document expose l'interopérabilité géosémantique, ses concepts, sa technologie et ses défis. La première section traite de l'origine et de l'utilité de l'interopérabilité géosémantique, la deuxième porte sur les principes fondamentaux et les exemples, la troisième décrit l'environnement technologique et la quatrième résume une étude de cas. Le document se termine par une exposition des défis et une conclusion.

**GéoConnexions** est une initiative de partenariat national dirigée par Ressources naturelles Canada. Elle facilite l'intégration et l'utilisation de l'**Infrastructure canadienne de données géospatiales (ICDG)**.

L'**ICDG** est une ressource en ligne qui permet d'améliorer l'échange, l'accessibilité et l'utilisation de l'information géospatiale – l'information sur des lieux géographiques du Canada. Elle aide les décideurs de tous les paliers de gouvernement, du secteur privé, des organismes non gouvernementaux et du milieu universitaire à prendre de meilleures décisions sur les priorités sociales, économiques et environnementales.

## 2. Introduction

### 2.1 Origines de l'interopérabilité géosémantique

Depuis des millénaires, les humains communiquent les uns avec les autres. La communication est passée des sons et des gestes simples à des expressions articulées, puis aux langues parlées et aux langues écrites. La communication est primordiale pour l'interaction humaine et donc pour interopérer. Pour ce faire, on se sert des langues pour encoder les messages sous forme de signes que le récepteur perçoit. Afin de pouvoir comprendre les signes, le récepteur doit utiliser un langage commun avec l'émetteur ou recourir à un traducteur. Au cœur d'une interopérabilité efficace se trouve le mode de transfert sémantique – c'est-à-dire la signification voulue des signes employés. L'interopérabilité géosémantique traite de cette question pour les données géospatiales transmises sur le Web. Pour ce faire, elle s'appuie sur des normes ainsi que sur des ontologies et des langages formels. En favorisant ces éléments, les infrastructures de données spatiales (IDS) facilitent l'interopérabilité géosémantique.

## 2.2 L'interopérabilité géosémantique dans la société d'aujourd'hui

---

Depuis cinq ans, le nombre des appareils intelligents reliés à l'Internet (voir l'Internet des choses) a dépassé le nombre d'êtres humains sur la terre et il continue d'augmenter rapidement. Des millions de personnes deviennent des producteurs de données (réseaux sociaux, sources d'information populaires). Les gouvernements adoptent des politiques d'accès libre aux données. Les villes, hyperconnectées, deviennent intelligentes, grâce aux données en temps réel, l'intelligence et les analyses géodécisionnelles. En technologies de l'information, les mégadonnées sont devenues le terme à la mode, les données géospatiales deviennent omniprésentes et le Web sémantique remplace progressivement le Web centré sur les documents. On s'approche du moment où les données recueillies en une seule journée dépasseront les données recueillies pendant toute l'histoire de l'humanité. S'il faut en croire Hoskins (cf. Bertolucci, 2013), « le matériel informatique a dominé pendant 20 ou 30 ans, les logiciels pendant 20 ou 30 ans », et on entre maintenant dans l'ère des données. Au seuil de cette nouvelle ère, les données géospatiales deviennent primordiales.

Pour éviter la confusion et les mauvaises interprétations, les systèmes du Web doivent se comprendre les uns les autres. Ils doivent être mis en interopérabilité sémantique. Selon Kuhn (2005), « l'interopérabilité sémantique est la seule forme d'interopérabilité utile ». Une interopérabilité sémantique de grande qualité est nécessaire pour économiser temps et argent lors de la recherche et de l'accès aux données, et pour l'évaluation de leur qualité pour une fin donnée, leur réutilisation, leur intégration et leur analyse, la découverte et l'évaluation des services Web, etc. Une interopérabilité géosémantique de grande qualité est donc essentielle en matière de données géospatiales.

## 3. Définitions de base de l'interopérabilité géosémantique

L'interopérabilité géosémantique entre les applications géospatiales sur le Web présente une analogie avec la communication et la collaboration humaine (Brodeur, 2012; Brodeur et collab., 2003; Kuhn, 2005). Par conséquent, il existe un risque de mauvaise interprétation (Sboui et Bédard, 2012). Pour comprendre le processus d'interopérabilité géosémantique, ses limites et ses solutions. Il faut disposer d'explications simples sur la manière dont on fait abstraction du monde pour communiquer dans un environnement basé sur le Web et sur les concepts, modèles, symboles, langages et la sémantique, sur les ontologies, l'interopérabilité, les contextes et la géosémantique. Les relations entre les sources classiques de définition, du type dictionnaires et thésaurus, sont expliquées, de même que les relations entre les normes particulières à chaque domaine.

### 3.1 Quels sont les « concepts » et comment les communiquer à des fins d'interopérabilité?

---

Le monde réel est complexe. Pour le comprendre, les humains créent des modèles cognitifs qui sont des vues abstraites de la réalité, comme l'image mentale d'une maison. Ces représentations abstraites ne retiennent que certains éléments caractéristiques : murs, toit, fenêtres, matériaux, dimensions, implantation sur la parcelle, proximité d'une route, mitoyenneté, accès, espaces de vie, sortie. Ce processus d'abstraction commence par la catégorisation de phénomènes présentant des caractéristiques et des comportements communs. Les maisons, parcelles, propriétaires et routes peuvent alors être considérés comme quatre catégories distinctes. Pour chacune d'elles, on se souviendra des caractéristiques et comportements présentant un intérêt afin de différencier les phénomènes individuels du groupe : l'adresse de la maison, la couleur, la localisation, le type de toit, les durées d'inoccupation. Il faut également se souvenir des relations qu'elles ont avec les autres membres de la même catégorie (p. ex., la maison M 1 est voisine de la maison M 2) ou avec les membres d'autres catégories (p. ex., la maison M1 se trouve le long de la route R1). Tous ces éléments abstraits sont appelés « concepts ».

Les concepts ne sont pas des phénomènes proprement dits, ils sont des substituts créés par le cerveau humain. Ils permettent de comprendre le monde, de s'en souvenir et de communiquer. Les ordinateurs n'ont pas de cerveau semblable au cerveau humain; cependant, leur structure et leur fonctionnement sont similaires à ce dernier et ils sont dotés de règles pour raisonner de manière automatique et se servant des concepts de leur modèle interne.

Pour communiquer et interopérer, les humains et les machines doivent encoder les concepts de leurs modèles internes propres sous une forme qui les rendra perceptibles et compréhensibles par d'autres humains ou machines. Le modèle physique ainsi obtenu consiste en symboles organisés (p. ex., des mots écrits, des sons) qui servent de substituts aux phénomènes de la réalité.

La triade « concept-symbole-phénomène » correspond à ce qu'on appelle le « triangle sémantique » : les concepts abstraits sont exprimés (la catégorie « maisons » du modèle cognitif) sous forme de symboles physiques (le mot « maison » écrit sur une page) qui renvoient aux phénomènes de la réalité (c'est-à-dire les maisons) qui sont abstraits sous forme de concepts dans le cerveau (la catégorie « maison »). Autrement dit, les humains et les ordinateurs ont recours à des symboles pour conceptualiser quelque chose dans le monde réel.

Le mode d'organisation des symboles est ce qu'on appelle la syntaxe. Les symboles et les règles syntaxiques constituent les éléments fondamentaux du langage (le français, les langages Java ou XML). Une personne qui parle plusieurs langues utilise différents symboles pour désigner le même phénomène, « house » ou « casa ». Si la personne ou la machine réceptrice d'un symbole emploie le même langage que l'émetteur, l'interprétation sera possible, la communication aura lieu et l'interopérabilité deviendra réalité.

### 3.2 Sources classiques de communication de la signification

---

Il existe différentes façons d'exprimer explicitement la signification acceptée des symboles utilisés. Les dictionnaires, les thésaurus, les lexiques et les glossaires attribuent une signification aux mots des langues naturelles (le français, par exemple). Toutefois, leurs définitions sont susceptibles de varier, d'où une

---

variabilité potentielle de leur correspondance avec les phénomènes du monde réel. À titre d'exemple, l'Oxford Dictionary en ligne définit le concept de « route » comme une « voie de communication importante qui mène d'un lieu à un autre, et dotée d'une surface spécialement préparée pour l'utilisation des véhicules » alors que la base de données lexicale en ligne WordNet définit le même concept comme « une voie ouverte, généralement publique, pour le déplacement et le transport ». Ces définitions d'ordre général s'avèrent souvent insuffisantes pour les applications géospatiales. Lors de la numérisation d'une route, par où commencer et par où finir lorsque cette route dessert plusieurs lieux? Que faire dans le cas d'une route non carrossable et fermée? S'agit-il encore d'une route? Devrait-on l'inclure dans la carte?

Les schémas sémantiques des bases de données, les dictionnaires de données et les catalogues de caractéristiques descriptives détaillent le sens, le contenu, le format et la structure des données aux fins des applications informatiques. Le Catalogue des entités de GéoBase au Canada définit une « route » comme « une section linéaire à la surface de la Terre qui a été conçue pour la circulation de véhicules ou qui en est le résultat. Un Élément routier est la représentation d'une route entre deux Jonctions. Un Élément routier est toujours limité par deux Jonctions. Un Élément routier est composé d'un ou plusieurs Segments routiers. » (Ressources naturelles Canada, 2012). Ce Catalogue d'entités comprend une description détaillée des types de données (un identifiant doit avoir 32 caractères), des codes (1 = même direction, 2 = direction contraire, 3 = non applicable), des instructions (la valeur « 0 » est utilisée quand aucune valeur ne s'applique, « 1 » est utilisé lorsque la valeur est inconnue), et des exemples (80-EST, 80-E, 80E). Ils contribuent tous à la compréhension de la signification des concepts. Un catalogue d'entités peut être détaillé jusqu'à un point suffisant pour servir à la spécification de l'acquisition des données (p. ex., la géométrie d'une route se situe en son centre, excepté pour les autoroutes qui exigent une ligne centrale pour chacune des voies).

Les normes de domaine proposent un vocabulaire commun et une organisation des données propres à faciliter les échanges entre systèmes. À titre d'exemple, la norme CGIS-NRN2 pour le Canada fournit les règles pour l'identification et la description des caractéristiques géographiques du réseau routier national. Plus ces normes sont universelles (p. ex., la norme S-101 de l'Organisation hydrographique internationale (OHI) en hydrographie), plus l'interopérabilité se répand dans le monde. L'absence de normes complique beaucoup la communication, tout comme l'absence de normes pour les panneaux de signalisation rendrait difficile l'interprétation de ces panneaux dans des territoires qui utilisent des langues différentes, parce que personne n'est en mesure d'interpréter correctement leur signification sans apprendre une nouvelle langue.

---

### **3.3 Qu'est-ce que la « sémantique » et quel est son rôle dans un environnement Web?**

---

La sémantique est ce qui permet de comprendre la signification des symboles et leur organisation (syntaxe). Elle analyse les rapports entre le langage employé par les humains ou les machines et les phénomènes auxquels renvoie ce langage dans le monde réel. Prenons un exemple, le mot anglais « pipe » désigne généralement une structure cylindrique fabriquée pour transporter de l'eau, du gaz ou du pétrole le long d'une route ou en dessous. Néanmoins, il désigne également un instrument pour fumer du tabac ou les tubes cylindriques qui produisent des sons dans un orgue ou même une barre verticale en informatique. Le même mot « pipe » en français est uniquement associé avec l'instrument du fumeur, alors qu'il n'a aucune signification en italien ou en espagnol. Par conséquent, un mot ou un symbole peut avoir plusieurs significations ou aucune, selon la langue utilisée. Le contraire est également vrai, plusieurs

---

symboles peuvent désigner le même phénomène. Pour désigner la structure cylindrique susmentionnée, on parlera de « conduit », « duct », « tuyau », « conduite » et de « conducto ».

Les applications informatiques obligent à donner un cadre formel à la sémantique pour minimiser les risques de mauvaise interprétation. La sémantique formelle emploie des langages mathématisés pour maîtriser les relations entre les symboles et les significations. Le nombre de concepts est réduit et le résultat est donc souvent spécifique à une application ou à un domaine, car il correspond à une convention acceptée entre les participants. On parle donc couramment de trois niveaux de vocabulaire (Guarino, 1998) : spécifique aux applications (les bases de données sur les rues), spécifique au domaine (la norme CGIS-NRN2) et mondiale (WordNet). Le premier niveau possède une sémantique plus spécifique, le dernier présente un risque supérieur de mauvaise interprétation.

Dans le Web sémantique, on considère qu'un « système possède un caractère sémantique s'il peut être traité et compris par un ordinateur » (Almeida et collab., 2011). Cela permet les échanges et les réutilisations de contenu (c'est-à-dire de données) entre les applications ou les entreprises et les communautés d'utilisateurs (World Wide Web Consortium [W3C], 2013). La sémantique est nécessaire pour trouver des données pertinentes.

### **3.4 Qu'est-ce qu'une « ontologie formelle », quels sont ses rôles et ses langages?**

---

La codification des ententes sémantiques formelles dans un environnement Web conduit généralement à la création d'ontologies fondées sur le langage informatique. En informatique, les ontologies sont des mécanismes explicites de spécification qui transposent les concepts en un langage lisible par ordinateur (Gruber, 1993; Guarino, 1998). Une ontologie formelle est « une représentation formelle des phénomènes à l'aide d'un vocabulaire sous-jacent assorti de définitions et d'axiomes qui rendent explicite la signification voulue et décrivent les phénomènes et leurs interrelations » (Brodeur et coll, 2003). Les éléments constitutifs d'une ontologie sont les concepts, relations et axiomes (Agarwal, 2005). La constitution d'une ontologie formelle représente un premier pas vers l'interopérabilité sémantique sur le Web (Kuhn, 2001). Le raisonnement automatique permet aux machines de découvrir des données, d'évaluer et de faire correspondre des ontologies similaires, d'intégrer et de transformer adéquatement les données provenant de multiples systèmes, de découvrir et de composer des services Web, etc.

Afin de définir les ontologies formelles pour le Web, on a mis au point des représentations graphiques et des langages. Pour faciliter l'élaboration d'ontologies, il existe des outils, notamment, Protégé (Université Stanford), Semaphore Ontology Manager (Smartlogic) et Vitro (Université Cornell).

### **3.5 Qu'est-ce que l'« interopérabilité »?**

---

L'interopérabilité est la capacité qu'ont les systèmes d'entrer en collaboration (d'interopérer). Elle permet les échanges de données et de services sur le Web. À titre d'illustration, si quelqu'un veut développer une nouvelle application pour la livraison de pizzas dans la province de Québec, le système pourrait détecter automatiquement le réseau routier des services Web de Ressources naturelles Canada, les adresses des maisons individuelles à partir du service Web de Géolocalisation Québec, et les données sur la circulation du service Web Transports Québec. L'utilisation des ontologies formelles intègre automatiquement ces données et applique des algorithmes de routage.

---



L'interopérabilité se situe à deux niveaux : 1) niveau technique (protocoles de communication, formats de données, langages informatiques, 2) niveau sémantique (interprétation de la signification des données pour l'obtention de résultats utiles et précis). Les organismes de normalisation s'attachent à l'un de ces niveaux (le « comment ») ou à l'autre (le « quoi ») ou encore aux deux. Toutefois, sans le dernier, le premier n'a aucune utilité.

### **3.6 Qu'est-ce que le « contexte »?**

---

La signification des symboles est toujours reliée à un contexte quelconque. Pour reprendre l'exemple précédent, quelqu'un qui consulte une carte de la municipalité de Winnipeg peut déduire que le mot « pipe » désigne de toute évidence la structure cylindrique qui se trouve sous la route ou à côté pour transporter l'eau, le gaz ou le pétrole. Le contexte indique qu'il s'agit d'une carte d'infrastructures municipales et que la langue utilisée à Winnipeg est l'anglais, ce qui conduit à éliminer toutes les autres significations potentielles de ce mot. « Le contexte fournit l'information pertinente liée aux circonstances dans lesquelles les données ont été définies et peuvent être utilisées » (Sboui et Bédard, 2012). Certains vont jusqu'à dire que le contexte est un ensemble de métadonnées organisées en ontologie. Par conséquent, les données de contexte aident à évaluer la qualité de l'interopérabilité sémantique. L'interopérabilité en contexte réduit les risques que les données soient mal interprétées et mal exploitées.

### **3.7 Qu'est-ce que la « géosémantique »?**

---

La géosémantique est la sémantique des phénomènes géographiques, ce qui comprend également la sémantique de leur géométrie. La sémantique de la géométrie implique deux composantes : la définition du type de donnée pour l'élément géométrique primitif (p. ex., ligne), la définition de la signification de la géométrie (p. ex., ligne centrale d'une route).

La géosémantique est absolument indispensable aux systèmes géospatiaux, car sans la sémantique de la géométrie, il est probable que plusieurs problèmes d'interopérabilité se posent. L'exemple le plus courant concerne les maisons sur les cartes à grande échelle. Dans de tels ensembles de données, un seul polygone symbolise une maison. Toutefois, ce polygone peut représenter le toit de la maison si les mesures ont été prises par photogrammétrie, ou ses fondations si les mesures ont été relevées par un arpenteur. Dans de tels cas, la signification même du terme « polygone de maison » diffère. Qui plus est, si les spécifications relatives à l'acquisition de données exigent d'inclure tous les détails d'une maison dépassant un mètre (balcons et entrées), la signification de ces polygones de maisons est différente de ceux où de tels détails n'ont pas à être mesurés.

La signification de primitive géométrique dans divers ensembles de données a également des retombées en matière d'interopérabilité géosémantique. Dans un ensemble de données, un « polygone de maison » désignera une seule surface avec son périmètre, alors que dans un autre ensemble de données, il pourra représenter une ligne fermée sans intérieur, ou même assimiler comme polygone unique l'agrégation de sous-polygones adjacents (en considérant comme un seul polygone une série de propriétés individuelles dans des maisons jumelées ou un condominium). La définition des primitives géométriques diffère habituellement d'un ensemble logiciel à l'autre, selon qu'il s'agit de systèmes d'information géographiques (SIG) ou de logiciels de conception assistée par ordinateur (CAO). L'Infrastructure canadienne de données géospatiales recommande fortement de recourir à des normes de géomatique

---

(OGC, ISO/TC-211, etc.) afin d'aplanir les difficultés liées à cet aspect de l'interopérabilité géosémantique.

Les différences dans la sémantique de la géométrie et la signification des primitives géométriques existent non seulement entre les divers systèmes, mais également entre les différentes époques d'un même système, en raison de l'évolution des spécifications au fil du temps en fonction de l'évolution des besoins. De telles différences, si elles passent inaperçues, conduiront à des erreurs dans la recherche automatique des meilleures sources de données sur les maisons, l'évaluation des correspondances entre les ontologies géospatiales, l'intégration et la transformation des données géométriques, etc. Elles aboutiront aussi à des analyses erronées ou incohérentes en ce qui concerne le nombre de maisons, la distance d'une maison à la ligne de rue, la densité d'occupation par mètre carré d'habitation. Trop souvent, on part du principe que l'interopérabilité sémantique est acceptable pour les données géospatiales et qu'il suffit d'utiliser un système de références spatiales commun. Il s'agit là d'une hypothèse erronée et il ne faut jamais oublier l'interopérabilité géosémantique en matière de données géospatiales.

### **3.8 Qu'est-ce que l'« interopérabilité géosémantique »?**

---

C'est la capacité des systèmes à employer des données et des systèmes géospatiaux de façon coopérative (en interopération) aux niveaux sémantiques et géométriques. L'interopérabilité géosémantique permet l'échange et l'utilisation en toute compatibilité sémantique des données et des services géométriques et non géométriques sur le Web. Elle interprète automatiquement et correctement la sémantique des données géométriques de même que le type de primitives géométriques pour produire des résultats précis et utiles. C'est la forme la plus riche d'interopérabilité pour les services et les données géospatiales.

### **3.9 Langages de l'interopérabilité géosémantique pour le Web**

---

Pour qu'il y ait interopérabilité géosémantique sur le Web, il faut que les humains expriment leurs concepts à l'aide de langages utilisables par les ordinateurs. Le Web sémantique fournit deux éléments pour traiter des aspects techniques et sémantiques de l'interopérabilité : 1) des formats communs pour l'intégration et l'association des données provenant de diverses sources, 2) un langage, le Resource Description Framework (RDF) pour enregistrer la relation entre les données et les objets du monde réel et les autres données. Le RDF fait le lien entre la syntaxe abstraite du Web et un cadre formel de sémantique et d'ontologie. C'est un équivalent basé sur des classes d'objets ou des diagrammes entités-relations (Wikipédia, 2014). Il utilise des triplets complexes : « sujet-prédicat-objet », similaires à « entité-attribut-valeur » pour associer explicitement les données (objet) aux aspects spécifiques (prédicat) d'une ressource Web (sujet). Un triplet RDF utilise une longue chaîne de caractères pour identifier de façon univoque les ressources Web du sujet et du prédicat grâce des identificateurs de ressources uniformes (URI). Ces URI comprennent généralement les adresses URL. Par conséquent, ce modèle simple de triplet permet aux machines de commencer la recherche de données dans une ressource basée sur le Web, puis de progresser dans un ensemble de bases de données ou de ressources reliées par leur contenu et leur sémantique, afin de trouver d'autres données connexes. Une telle collection d'ensembles de données interdépendants sur le Web est connue sous le terme de données liées (Consortium World Wide Web [W3C ], 2013). Il peut être vu comme une base de données mondiale susceptible d'être interrogée en utilisant SPARQL, un langage de requête servant à exprimer des requêtes dans les diverses sources de données RDF (World Wide Web Consortium, 2013) et GEOSPARQL, un langage de requêtage géographique RDF (Open Geospatial Consortium, 2012).

---

Des langages plus complexes que le RDF sont également utilisés pour l'interopérabilité. En utilisant des caractéristiques RDF, RDFS (schéma RDF) prend en charge la sémantique supplémentaire tandis que le Web Ontology Language (OWL) utilise les caractéristiques RDFS pour augmenter encore la puissance expressive. En matière de services, le WSDL (Web Service Description Language) décrit la syntaxe des services Web tandis que Web Ontology Language pour les services Web (OWL-S) est une ontologie qui soutient la spécification sémantique pour les services.

### **3.10 Normes de soutien de l'interopérabilité géosémantique**

---

Plusieurs organismes de normalisation ont mis au point des normes pour soutenir l'interopérabilité géosémantique. L'ISO/TC-211 publie de nombreuses normes concernant les modèles de données pour l'information géographique, sa gestion, les services d'information géographique et l'encodage de cette information (Organisation internationale de normalisation [ISO], 2009). L'Open Geospatial Consortium (OGC) développe des normes pour l'interopérabilité des services géospatiaux sur le Web (WMS, WFS, WCS, WPS), KML, LBG et GeoSPARQL. L'OGC propose également des normes pour les applications destinées aux collectivités (aviation, défense, urgences, installations). Le World Wide Web Consortium (W3C) propose plusieurs spécifications pour le Web sémantique et les données liées : plusieurs séries de normes, notamment, RDF, OWL, SPARQL, WSDL, RIF et SKOS. Nombre de domaines fortement réglementés disposent de normes spécialisées, entre autres, l'hydrographie (Organisation hydrographique internationale, 2008), le transport (cf. ISO/TC 204 de l'Organisation internationale de normalisation, 2014) et l'aviation (AIXM, Eurocontrol ou Organisation européenne pour la sécurité de la navigation aérienne, 2014).

## **4. Étude de cas : Ordnance Survey**

Ordnance Survey (OS) est l'organisme cartographique national en Grande-Bretagne. Cet organisme est reconnu pour ses travaux d'avant-garde dans le domaine de l'interopérabilité géosémantique. L'étude de cas présentée est fondée sur une entrevue avec M. Glen Hart et M. John Goodwin et sur la documentation qu'ils ont fournie.

Ordnance Survey a commencé, en 2004, par s'intéresser aux ontologies et à l'interopérabilité pour faciliter l'intégration des données. Elle a alors testé des stratégies et des outils divers pour construire des ontologies jusqu'en 2006. En 2007-2008, OS a constitué sa première ontologie, enrichissant ainsi sa connaissance des données afin d'offrir de nouveaux contrôles de la qualité. En 2009, OS est parvenue à convaincre les entreprises et les autorités, dans le contexte de l'initiative sur les données libres du Royaume-Uni, de publier leurs données sous forme de données liées. En 2010, le premier ensemble de données, présenté sous le format des données liées (c'est-à-dire RDF) concernait les limites d'OS, autrement dit, les limites de comtés et de districts. OS a construit une ontologie pour les relations spatiales en incluant explicitement les relations topologiques pour faciliter les requêtes spatiales du type : contenant, à l'intérieur, touchant. Le premier ensemble de données ne comprenait pas de coordonnées afin de préserver leur valeur commerciale. À la suite de ce lancement, l'équipe a travaillé sur une plateforme pour publier les données liées. OS a alors lancé trois ensembles de données liées : OS Boundary Line (les limites administratives), Code point (les unités de codes postaux) et 50K Gazetteer (la liste des noms

---

géographiques provenant d'une carte au 1/50 000, balisée de points représentant des couches cartographiques de 1 km<sup>2</sup>). De plus, ces trois ensembles de données publiées sont maintenant offerts dans un ensemble associant les données liées d'Ordnance Survey. Les réactions positives se sont manifestées sous forme d'un accroissement continu de l'utilisation de ces ensembles de données, pour faire le lien entre les données liées des limites administratives avec les données statistiques des autres organismes gouvernementaux.

OS a alors développé l'application « See UK » (« Voir le R.-U. » sur <http://apps.seme4.com/see-uk/>) pour fournir diverses données liées, concernant la criminalité, les écoles, les transports, en cliquant sur une carte à divers niveaux. Deux entreprises, Talis et Iconomical, et l'équipe de data.gov.uk ont mis au point le Research Funding Explorer pour démontrer l'intérêt réel des techniques des données liées en ce qui concerne l'assemblage et l'analyse des données provenant d'un certain nombre de sources différentes. Cette application permet à l'utilisateur de consulter les projets de recherche financés par sujet, par organisation ou par lieu géographique grâce à une liaison avec les données liées du point OS Code en remplaçant le texte du code postal par l'URI du code postal.

Selon OS, l'utilisation des données liées présente les avantages suivants : souplesse, adaptabilité et intégration simplifiée des données en comparaison des structures rigides des bases de données et des divers formats de données. De tels avantages permettent d'éliminer les cloisonnements des silos de données et de détecter les problèmes dans les données. OS travaille dans le but d'affecter un URI à tous les « lieux » de Grande-Bretagne et ses projets futurs porteront sur la gestion de la qualité, la fiabilité et la provenance des données, la concession de licences, le contrôle des versions, le raisonnement et la découverte automatique des données, la conversion et l'importation à partir du Web et des capteurs.

## 5. Défis

L'interopérabilité des écosystèmes géospatiaux se heurte à la difficulté de convaincre les organismes d'investir les efforts et les compétences nécessaires pour bien décrire les niveaux sémantique et géométrique de leurs données. En l'absence de tels efforts, la version formelle en code machine de la signification des données qui existe dans l'esprit des organismes risque d'être mal interprétée. Cette constatation vaut particulièrement là où les contextes diffèrent (époques, régions, territoires, domaines, langues naturelles, cultures, etc.). Par ailleurs, plusieurs phénomènes ont, par leur nature même, un caractère flou. Même si les humains peuvent composer avec des concepts « relativement clairs », les machines ont absolument besoin de concepts formels. Le recours aux données de contexte est prometteur dans l'évaluation de l'adaptation de l'interopérabilité géosémantique à des usages déterminés et pour alerter automatiquement les usagers des risques de mauvaise interprétation, comme le suggèrent Sboui et Bédard (2012). L'inclusion de cette prise en compte du contexte dans un processus automatique est la solution pour une interopérabilité fiable et constructive.

Sur le plan technique, Goodwin et Hart de l'OS ont recensé deux éléments primordiaux : 1) l'expression du temps et des métadonnées pertinentes de manière utilisable dans les données liées, 2) l'inclusion des données adéquates du contexte pour réduire les erreurs. Ils ont également noté l'absence actuelle de normes et de technologies à maturité pour le Web sémantique.

---

## 6. Conclusion

L'interopérabilité géosémantique concerne les aspects tant sémantiques que géométriques de l'information géospatiale. Les aspects géométriques incluent le type de primitives géométriques et la signification fondée sur ses spécifications d'acquisition. Les contextes entourant les données peuvent varier grandement de façon à rendre l'interopérabilité géosémantique inapplicable pour telle ou telle utilisation. L'évaluation automatique de la qualité de l'interopérabilité géosémantique est réalisable à l'aide de l'information sur le contexte. L'interopérabilité géosémantique de qualité est la forme la plus riche d'interopérabilité pour les données et services géospatiaux. L'interopérabilité géosémantique de qualité est le pas qui doit nécessairement être franchi dans la voie menant au Web géosémantique. La création d'un tel Web est nécessaire aux écosystèmes technologiques de demain que nous apportent l'Internet des choses (IoT), les villes intelligentes et la société géomatisée.

# Annexe 1 : Liste d'acronymes

Tableau de la définition des acronymes utilisés dans le présent document.

Acronyme/abréviation	Définition
AIXM	Modèle d'échange d'informations aéronautiques (Aeronautical Information Exchange Model)
CGIS-NRN2	Norme canadienne de l'information géographique - Réseau routier national 2.0
DAML+OIL	DARPA Agent Modeling Language (langage de modélisation à base d'agents) + Ontology Inference Layer (couche d'inférence d'ontologie)
DARPA	Defense Advanced Research Projects Agency (Agence pour les projets de recherche avancée de défense)
FME	Feature Manipulation Engine (outil d'extraction des données géographiques)
HTML	HyperText Markup Language (langage de balisage hypertexte)
IPA	Interface de programmation d'applications
JSON-LD	JavaScriptObject Notation (Notation d'objet JavaScript) – données liées
KML	Keyhole Markup Language (langage à base de balises géolocales)
LBG	Langage de balisage géographique
Norme ISO/TC	Organisation internationale de normalisation/Comité technique
OGC	Open Geospatial Consortium (consortium pour développer des standards ouverts)
OHI	Organisation hydrographique internationale
OIL	Ontology Inference Layer (couche d'inférence d'ontologie )
OS	Ordnance Survey (Organisation nationale de cartographie de la Grande-Bretagne)
OWL	Ontology Web Language (langage ontologique Web )
OWLGrEd	Ontology Web Language Graphical Editor (éditeur graphique de langage ontologique Web)
OWL-S	Ontology Web Language (langage ontologique Web) – Service
RAGLD	Rapid Assembly of Geo-centred Linked Data (assemblage rapide des données géographiques liées)
RDF	Resource Description Framework (cadre de description des ressources)
RDFa	Resource Description Framework in Attributes (cadre de description des ressources dans attributs)
RDFS	Resource Description Framework Schema (schéma du cadre de description des ressources)
RIF	Rule Interchange Format (format d'échange de règles)
SCW	Service de cartes Web
SIA	Services d'information aéronautique
SIG	Système d'information géographique

Acronyme/abréviation	Définition
SKOS	Simple Knowledge Organization System (système simple d'organisation des connaissances)
TI	Technologie de l'information
URI	Uniform Resource Identifier (identificateur de ressources uniforme)
W3C	World Wide Web Consortium
WCS	Web Coverage Service (service de couverture pour le Web)
WFS	Web Feature Service (service d'entités Web)
WPS	Web Processing Service (service de traitement Web)
WSDL	Web Services Description Language (langage de description des services Web)
XML	Extensible Markup Language (langage de balisage extensible)

## Annexe 2 : Références

Agarwal, P., 2005. « Ontological considerations in GIScience ». *International Journal of Geographical Information Science*, 19(5), p. 501-536.

Almeida, M.B., R.R. Souza, F. Fonseca (2011). Semantics in the Semantic Web: a Critical Evaluation. *Knowledge Organization Journal*, vol. 38, No 3, p. 187-203

Bertolucci, J. (4 janvier 2013). « The Age Of 'Data Ubiquity': Sensors Spread ». Extrait de *InformationWeek*: <http://www.informationweek.com/big-data/big-data-analytics/the-age-of-data-ubiquity-sensors-spread/d/d-id/1109327?>

Brodeur, J. (2012). « Geosemantic Interoperability and the Geospatial Semantic Web » Dans Springer (ed), *Handbook of Geographic Information*, Chap. 15, p. 291-310

Brodeur, J., Bédard, Y., Edwards, G. et Moulin, B. (2003). « Revisiting the Concept of Geospatial Data Interoperability within the Scope of Human Communication Processes ». *Transactions in GIS*, 7 (2), p. 243-265.

Eurocontrol. (2014). « AIXM Aeronautical Information Exchange Model », Extrait de Eurocontrol: [http://www.aixm.aero/public/subsite\\_homepage/homepage.html](http://www.aixm.aero/public/subsite_homepage/homepage.html)

Gruber, T.R. (1993). *A Translation Approach to Portable Ontology Specification*. KSL 92-71 (Knowledge Systems Laboratory, Stanford)

Guarino, N. (1998). *Formal ontology and information systems. Proceedings, Formal Ontologies Information Systems (FOIS '98)*, Trento, p. 3–15

IEEE. (1991). *610-1991 - IEEE Standard Computer Dictionary. A Compilation of IEEE Standard Computer Glossaries*. New York: IEEE.



Organisation hydrographique internationale (2008). *Normes OHI pour les levés hydrographiques*.  
Organisation hydrographique internationale

Organisation internationale de normalisation (ISO). (2014). *ISO TC 204 Normes systèmes intelligents de transports*. Extrait d'ISO : [http://www.iso.org/iso/iso\\_technical\\_committee?commid=54706](http://www.iso.org/iso/iso_technical_committee?commid=54706)

Organisation internationale de normalisation (ISO). (2009). *ISO TC 211 Standards Guide*. ISO.

Powell, J. (June 2011). « IHO S-101, the Next Generation ENC Product Specification ». *TransNav International Journal on Marine Navigation*. Vol. 5, No. 2, p. 167-171.

Kuhn, W. (2001). « Ontologies In Support Of Activities In Geographical Space ». *International Journal of Geographical Information Science*, vol. 15, No 17, p. 613-631.

Kuhn, W. (2005). « Geospatial Semantics: Why, of What, and How? » *Journal on Data Semantics III*, 1-24

Ressources naturelles Canada, 2012. Réseau routier national, Catalogue d'entités Vue segmentée Édition 2.0.1. 34 p.

Onsrud, H.J. (2010). « Legal Interoperability in Support of Spatially Enabling Society ». Dans A. Rajabifard, J. Crompvoets, M. Kalantari et B. Kok, « Spatially Enabling Society: Research, Emerging Trends and Critical Assessment », p. 163-172. Ouvrage complet à l'adresse suivante [http://memberservices.gsdi.org/files/?artifact\\_id=902](http://memberservices.gsdi.org/files/?artifact_id=902)

Open Geospatial Consortium (2012). *GEOSPARQL A geographic Query Language for RDF Data*. OG.

Sboui, T., Y. Bédard (2012). « Universal Geospatial Ontology for the Semantic Interoperability of Data: What are the Risks and How to Approach them? » Dans Podobnikar, T. & Ceh, M. (ed(s)), *Universal Ontology of Geographic Space: Semantic Enrichment for Spatial Data*, Chap. 1, p. 1-27

Wikipédia. (11 novembre 2013). *RDF Schema*. Extrait de Wikipédia : [http://en.wikipedia.org/wiki/RDF\\_Schema](http://en.wikipedia.org/wiki/RDF_Schema)

Wikipédia. (9 février 2014). *Simple Knowledge Organization System*. Extrait de Wikipédia : <http://en.wikipedia.org/wiki/SKOS>

World Wide Web Consortium. (22 mars 2013). *RDF - Resource Description Framework*. Extrait de W3C : <http://www.w3.org/RDF/>

World Wide Web Consortium. (26 mars 2013). *SPARQL Query Language for RDF*. Extrait de W3C : <http://www.w3.org/TR/rdf-sparql-query/>

World Wide Web Consortium. (20 mai 2013). *Ontology Editors*. Extrait de W3C Wiki : [http://www.w3.org/wiki/Ontology\\_editors](http://www.w3.org/wiki/Ontology_editors)

World Wide Web Consortium. (19 juin 2013). *W3C Semantic Web Activity*. Extrait de W3C : <http://www.w3.org/2001/sw/>



World Wide Web Consortium. (11 décembre 2013) *OWL - Ontology Web Language*. Extrait de W3C : <http://www.w3.org/2001/sw/wiki/OWL>

World Wide Web Consortium. (2014). *RDF Primer 1,1*. W3C